Foundations of Informatics: a Bridging Course

Part C: Context-Free Languages
Lesson 4: The Word Problem for Context-Free Languages

Thomas Noll
Software Modeling and Verification Group
RWTH Aachen University

https://moves.rwth-aachen.de/people/noll/

Software Modeling and Verification Chair

RWTH AACHEN UNIVERSITY

The Word Problem for CFL

Word Problem for ε-free CFL

Given CFG $G = \langle N, \Sigma, P, S \rangle$ such that $\varepsilon \notin L(G)$ and $w \in \Sigma^+$, decide whether $w \in L(G)$ or not.

(If $w = \varepsilon$, then $w \in L(G)$ easily decidable for arbitrary G)

The Word Problem for CFL

Word Problem for ε-free CFL

Given CFG $G = \langle N, \Sigma, P, S \rangle$ such that $\varepsilon \notin L(G)$ and $w \in \Sigma^+$, decide whether $w \in L(G)$ or not.

(If $w = \varepsilon$, then $w \in L(G)$ easily decidable for arbitrary G)

Algorithm (by Cocke, Younger, Kasami – CYK algorithm)

1. *Transform G into Chomsky NF*
2. *Let $w = a_1 \dots a_n$ ($n \geq 1$)*
3. *Let $w[i,j] := a_i \dots a_j$ for every $1 \leq i \leq j \leq n$*
4. *Consider segments $w[i,j]$ in order of increasing length, starting with $w[i,i] = a_i$ (i.e., letters)*
5. *In each case, determine $N_{i,j} := \{A \in N \mid A \Rightarrow^* w[i,j]\}$ using a "dynamic programming" approach:*
   - *$i = j$: $N_{i,i} = \{A \in N \mid A \to a_i \in P\}$*
   - *$i < j$: $N_{i,j} = \{A \in N \mid \exists B, C \in N, k \in \{i, \dots, j-1\} : A \to BC \in P, B \in N_{i,k}, C \in N_{k+1,j}\}$*
6. *Test whether $S \in N_{1,n}$ (and thus, whether $S \Rightarrow^* w[1,n] = w$)*

Matrix Representation of CYK Algorithm

	a_1	a_2	a_3	$\cdots$	a_n
$i \backslash j$	1	2	3	$\cdots$	n
1	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	$\cdots$	$N_{1,n}$
2	X	$N_{2,2}$	$N_{2,3}$	$\cdots$	$N_{2,n}$
3	X	X	$N_{3,3}$	$\cdots$	$N_{3,n}$
$\vdots$	$\vdots$	$\vdots$		$\cdots$	$\vdots$
n	X	X	$\cdots$	$\cdots$	$N_{n,n}$

Matrix Representation of CYK Algorithm

$$
\begin{aligned}
N_{1,1} &= \{A \in N \mid A \to a_1 \in P\} \\
N_{2,2} &= \{A \in N \mid A \to a_2 \in P\} \\
&\vdots
\end{aligned}
$$

	a_1	a_2	a_3	$\cdots$	a_n
$i \backslash j$	1	2	3	$\cdots$	n
1	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	$\cdots$	$N_{1,n}$
2	X	$N_{2,2}$	$N_{2,3}$	$\cdots$	$N_{2,n}$
3	X	X	$N_{3,3}$	$\cdots$	$N_{3,n}$
$\vdots$	$\vdots$	$\vdots$		$\cdots$	$\vdots$
n	X	X	$\cdots$	$\cdots$	$N_{n,n}$

Matrix Representation of CYK Algorithm

	a_1	a_2	a_3	$\cdots$	a_n
$i \backslash j$	1	2	3	$\cdots$	n
1	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	$\cdots$	$N_{1,n}$
2	X	$N_{2,2}$	$N_{2,3}$	$\cdots$	$N_{2,n}$
3	X	X	$N_{3,3}$	$\cdots$	$N_{3,n}$
$\vdots$	$\vdots$	$\vdots$		$\cdots$	$\vdots$
n	X	X	$\cdots$	$\cdots$	$N_{n,n}$

$$
\begin{aligned}
N_{1,1} &= \{A \in N \mid A \to a_1 \in P\} \\
N_{2,2} &= \{A \in N \mid A \to a_2 \in P\} \\
&\vdots \\
N_{1,2} &= \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{1,1}, C \in N_{2,2}\} \\
N_{2,3} &= \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{2,2}, C \in N_{3,3}\} \\
&\vdots
\end{aligned}
$$

Matrix Representation of CYK Algorithm

	a_1	a_2	a_3	$\cdots$	a_n
$i \backslash j$	1	2	3	$\cdots$	n
1	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	$\cdots$	$N_{1,n}$
2	X	$N_{2,2}$	$N_{2,3}$	$\cdots$	$N_{2,n}$
3	X	X	$N_{3,3}$	$\cdots$	$N_{3,n}$
$\vdots$	$\vdots$	$\vdots$		$\cdots$	$\vdots$
n	X	X	$\cdots$	$\cdots$	$N_{n,n}$

$$
\begin{aligned}
N_{1,1} &= \{A \in N \mid A \to a_1 \in P\} \\
N_{2,2} &= \{A \in N \mid A \to a_2 \in P\} \\
&\vdots \\
N_{1,2} &= \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{1,1}, C \in N_{2,2}\} \\
N_{2,3} &= \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{2,2}, C \in N_{3,3}\} \\
&\vdots \\
N_{1,3} &= \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{1,1}, C \in N_{2,3}\} \\
&\cup \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{1,2}, C \in N_{3,3}\} \\
N_{2,4} &= \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{2,2}, C \in N_{3,4}\} \\
&\cup \{A \in N \mid \exists B, C \in N : A \to BC \in P, B \in N_{2,3}, C \in N_{4,4}\} \\
&\vdots
\end{aligned}
$$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1						
2	X					
3	X	X				
4	X	X	X			
5	X	X	X	X		
6	X	X	X	X	X	

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$					
2	X					
3	X	X	$\{S\}$			
4	X	X	X	$\{S\}$		
5	X	X	X	X		
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$					
2	X	$\{B\}$				
3	X	X	$\{S\}$			
4	X	X	X	$\{S\}$		
5	X	X	X	X	$\{B\}$	
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$				
2	X	$\{B\}$				
3	X	X	$\{S\}$	$\emptyset$		
4	X	X	X	$\{S\}$	$\emptyset$	
5	X	X	X	X	$\{B\}$	
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$				
2	X	$\{B\}$	$\{A\}$			
3	X	X	$\{S\}$	$\emptyset$		
4	X	X	X	$\{S\}$	$\emptyset$	
5	X	X	X	X	$\{B\}$	$\{A\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$				
2	X	$\{B\}$	$\{A, B\}$			
3	X	X	$\{S\}$	$\emptyset$		
4	X	X	X	$\{S\}$	$\emptyset$	
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$			
2	X	$\{B\}$	$\{A, B\}$			
3	X	X	$\{S\}$	$\emptyset$		
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$			
2	X	$\{B\}$	$\{A, B\}$	$\{A \quad\}$		
3	X	X	$\{S\}$	$\emptyset$		
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$			
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$		
3	X	X	$\{S\}$	$\emptyset$		
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$			
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$		
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$		
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$		
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$		
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$	$\{B\}$	
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$		
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$	$\{B\}$	
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	$\emptyset$
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$	$\emptyset$	
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$	$\{B\}$	
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	$\emptyset$
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$	$\emptyset$	
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$	$\{B\}$	$\{A \quad\}$
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	$\emptyset$
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$	$\emptyset$	
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$	$\{B\}$	$\{A, B\}$
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	$\emptyset$
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G: \; S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$	$\emptyset$	$\{S\}$
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$	$\{B\}$	$\{A, B\}$
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	$\emptyset$
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

Applying the CYK Algorithm

Example

- $G:\ S \to SA \mid a$
  $A \to BS$
  $B \to BB \mid BS \mid b \mid c$
- $w = abaaba$

	a	b	a	a	b	a
$i \backslash j$	1	2	3	4	5	6
1	$\{S\}$	$\emptyset$	$\{S\}$	$\{S\}$	$\emptyset$	$\{S\}$
2	X	$\{B\}$	$\{A, B\}$	$\{A, B\}$	$\{B\}$	$\{A, B\}$
3	X	X	$\{S\}$	$\emptyset$	$\emptyset$	$\emptyset$
4	X	X	X	$\{S\}$	$\emptyset$	$\{S\}$
5	X	X	X	X	$\{B\}$	$\{A, B\}$
6	X	X	X	X	X	$\{S\}$

$$S \in N_{1,6} \implies w = abaaba \in L(G)$$

Summary: The Word Problem for Context-Free Languages

Seen:

- Given CFG G and $w \in \Sigma^*$, decide whether $w \in L(G)$ or not
- Decidable using CYK algorithm (based on dynamic programming)
- Cubic complexity

Summary: The Word Problem for Context-Free Languages

Seen:

- Given CFG G and $w \in \Sigma^*$, decide whether $w \in L(G)$ or not
- Decidable using CYK algorithm (based on dynamic programming)
- Cubic complexity

Next:

- Emptiness problem